Near-equilibrium measurements of non-equilibrium free energy

David A. Sivak and Gavin E. Crooks

Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
(Dated: December 14, 2009)

A central endeavor of thermodynamics is the measurement of free energy changes. Regrettably, although we can measure the free energy of a system in thermodynamic equilibrium, typically all we can say about the free energy of a non-equilibrium ensemble is that it is larger than that of the same system at equilibrium. Herein, we derive a formally exact expression for the probability distribution of a driven system, which involves path ensemble averages of the work over trajectories of the time-reversed system. From this we find a simple near-equilibrium approximation for the free energy in terms of an excess mean time-reversed work, which can be experimentally measured on real systems. With analysis and computer simulation, we demonstrate the accuracy of our approximations for several simple models.

PACS numbers: 05.70.Ln,05.40.-a,89.70.Cf

Recent advances in nanotechnology make it increasingly possible to engineer molecular scale structures for the deliberate and efficient manipulation of energy, matter and information on the nanometer scale. Artificial microscopic machines include heat pumps designed for very localized cooling; osmotic membranes built from carbon nanotubes; quantum logic gates designed to manipulate and stabilize quantum information; nano-structured thermoelectrics; devices for the capture and separation of carbon dioxide; and efficient photovoltaic solar cells.

Notably, molecular scale machines typically operate far from thermodynamic equilibrium, limiting the applicability of equilibrium statistical mechanics. In particular, while at equilibrium the free energy of a system is minimized (given the external constraints) and is often relatively easily measured, out of equilibrium no standard measurement technique exists, impeding the quantitative understanding of non-equilibrium behavior. To partially redress this deficit, we herein develop an experimentally tractable approach to measure the free energy of systems away from equilibrium. We find that to a strikingly good approximation, the difference between the free energy of a non-equilibrium ensemble and the equivalent system in equilibrium is determined by an excess mean time-reversed work, Eq. (13).

We consider a physical system with a collection of controllable parameters λ, in contact with a constant temperature heat bath with reciprocal temperature $\beta = (k_{\rm B}T)^{-1}$, where $k_{\rm B}$ is Boltzmann's constant. The free energy F of the system, in or out of equilibrium, can be defined as $\langle E \rangle - S/\beta$, for mean energy $\langle E \rangle$ and entropy $S = -\sum_x P(x) \ln P(x)$ in natural units. Here, x labels the microstates of the system. Interestingly, the free energy difference between any non-equilibrium ensemble $P_\lambda^{\rm neq}$ and the corresponding canonical equilibrium ensemble $P_\lambda^{\rm eq}(x) = \exp\{\beta\left[F_\lambda^{\rm eq} - E_\lambda(x)\right]\}$ with the same control parameters λ is equal to the relative entropy between the two ensembles [1]:

$$
\begin{aligned}
D(P_\lambda^{\rm neq}\|P_\lambda^{\rm eq}) &= -S_\lambda^{\rm neq} - \sum_x P_\lambda^{\rm neq}(x)\,\beta\left[F_\lambda^{\rm eq} - E_\lambda(x)\right] \\
&= -S_\lambda^{\rm neq} - \beta F_\lambda^{\rm eq} + \beta\langle E_\lambda\rangle_{\rm neq} \\
&= \beta\left(F_\lambda^{\rm neq} - F_\lambda^{\rm eq}\right) \ .
\end{aligned}
\tag{1}
$$

Thus, in both a thermodynamic and information theoretic sense, the free energy difference between nonequilibrium and equilibrium ensembles measures a distance from equilibrium.

A non-equilibrium ensemble is specified by a protocol Λ that describes the history of the control parameters over some time interval. In the corresponding time-reversed protocol $\widetilde{\Lambda}$, the system starts with the final parameters $\lambda_{\rm b}$ of the forward protocol, and then the controllable parameters retrace the same series of changes, in reverse, to end with the initial values $\lambda_{\rm a}$ of the forward protocol. Measurements performed on a system using a pair of conjugate protocols Λ and $\widetilde{\Lambda}$ are related by [2],

$$
\langle \mathcal{A} \rangle_{\lambda_a;\Lambda} = \langle \widetilde{\mathcal{A}}\,e^{-\beta\mathcal{W}}\rangle_{\lambda_b;\widetilde{\Lambda}} \Big/ \langle e^{-\beta\mathcal{W}}\rangle_{\lambda_b;\widetilde{\Lambda}} \ .
\tag{2}
$$

Here, $\mathcal{A}$ is a measurement of the system (any real function of the phase space trajectory), $\widetilde{\mathcal{A}}$ is the corresponding time-reversed measurement, and $\mathcal{W}$ is the work performed on the system. The angled brackets indicate that measurements are averaged over an experimental protocol, specified by subscripts: the first subscript indicates the initial preparation of the system; the second subscript, after the semicolon, indicate the protocol during measurement. Thus "$\lambda_{\rm a};\Lambda$" specifies that the system is equilibrated with fixed parameters $\lambda_{\rm a}$ and then the properties of the system are measured while the system is driven with protocol Λ, whereas "$\lambda_{\rm b};\widetilde{\Lambda}$" indicates initial equilibration at $\lambda_{\rm b}$ followed by measurement during the time-reversed protocol $\widetilde{\Lambda}$. If the preparation protocol is not explicitly stated, as is the case in many of our previous papers, then implicitly the system is prepared

at equilibrium with the initial control parameters of the measurement protocol.

We will use this expression to relate non-equilibrium probability distributions to cumulants of work. First, we replace the generic measurement $\mathcal{A}$ with a delta function of the final system configuration $\delta\left(x - x_{\mathrm{b}}\right)$. This gives a relation between the non-equilibrium probability of a configuration, and a non-linear average of the work performed on the system during the time-reversed protocol, starting from that configuration [2]:

$$
\begin{aligned}
P_\Lambda(x) &= \left\langle \delta\left(x - x_{\mathrm{b}}\right) \right\rangle_{\lambda_{\mathrm{a}};\Lambda} \\
&= \left\langle \delta(x - x_{\mathrm{b}})e^{-\beta\mathcal{W}} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \Big/ \left\langle e^{-\beta\mathcal{W}} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \\
&= P_{\lambda_{\mathrm{b}}}(x) \left\langle e^{-\beta\mathcal{W}} \right\rangle_{x;\widetilde{\Lambda}} \Big/ \left\langle e^{-\beta\mathcal{W}} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \ . \quad (3)
\end{aligned}
$$

The subscript "$x;\widetilde{\Lambda}$" indicates initial preparation of the system in microstate x and subsequent work measurement during protocol $\widetilde{\Lambda}$. Next, we rearrange the previous expression as in Ref. [3],

$$
\ln \frac{P_\Lambda(x)}{P_{\lambda_{\mathrm{b}}}(x)} = \ln \frac{\left\langle e^{-\beta\mathcal{W}} \right\rangle_{x;\widetilde{\Lambda}}}{\left\langle e^{-\beta\mathcal{W}} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}} \ , \quad (4)
$$

and expand both non-equilibrium averages as a sum of cumulants $\kappa^{(n)}$ using the generating function, $\ln\left\langle e^{cy} \right\rangle = \sum_{n=1}^\infty \frac{c^n}{n!}\kappa^{(n)}(y)$ [4], giving

$$
\ln \frac{P_\Lambda(x)}{P_{\lambda_{\mathrm{b}}}(x)} = -\beta\left(\left\langle \mathcal{W} \right\rangle_{x;\widetilde{\Lambda}} - \left\langle \mathcal{W} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \right) + K_{x;\widetilde{\Lambda}} \quad (5)
$$

where

$$
K_{x;\widetilde{\Lambda}} = \sum_{n=2}^\infty \frac{(-\beta)^n}{n!} \left[\kappa^{(n)}_{x;\Lambda}(W) - \kappa^{(n)}_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}(W) \right] \ . \quad (6)
$$

Recall that the first cumulant is the mean and the second the variance. Each term in the expansion is an 'excess' work cumulant, the difference between work cumulants of the time-reversed protocol, one when the system is initially prepared in a particular microstate x and one when initially prepared in an equilibrium-weighted ensemble of initial microstates. Given that the phase space distributions relax to the same steady state, irrespective of the initial microstate, it follows that each excess cumulant is finite, and converges to its asymptotic value in the long time limit [5]. Note that our excess mean work differs substantially from the excess heat of Oono and Paniconi [6]: ours is a time-reversed quantity, is defined over a finite time rather than instantaneously, and is an excess over a different baseline.

These relations for non-equilibrium probabilities are formally exact, yet impractical. In particular, if a system is driven into a non-equilibrium steady state by a periodic perturbation, then the work performed on the system increases without bound. Both the exponential average in

Eq. (4) and the higher-order cumulants in Eq. (6) are dominated by low dissipation realizations of the protocol, which are extremely rare and not observable in any realistic experiment [7].

To proceed further, we require a tractable approximation to the exact non-equilibrium probability. Truncating the cumulant expansion at the first term is tempting [8], but in practice the higher-order cumulants cannot be entirely neglected. We instead make a weaker assumption, that the higher-order excess cumulants of the work are *uninformative* about the non-equilibrium probability distributions, and therefore that the sum of higher-order excess work cumulants is a constant, independent of initial microstate:

$$
K_{x;\widetilde{\Lambda}} \equiv K_{\widetilde{\Lambda}} \ . \quad (7)
$$

We next develop an approximation for this constant $K_{\widetilde{\Lambda}}$. Consider the relative entropy $D(p\|q) = \sum_x p_x \ln(p_x/q_x)$ between the non-equilibrium and corresponding equilibrium ensembles, under our first approximation [Eq. (7)]:

$$
D(P_\Lambda \| P_{\lambda_{\mathrm{b}}}) \approx -\beta\left(\left\langle \mathcal{W} \right\rangle_{\Lambda;\widetilde{\Lambda}} - \left\langle \mathcal{W} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \right) + K_{\widetilde{\Lambda}} \quad (8)
$$

$$
D(P_{\lambda_{\mathrm{b}}} \| P_\Lambda) \approx -K_{\widetilde{\Lambda}} \ . \quad (9)
$$

Here "$\Lambda;\widetilde{\Lambda}$" denotes initial preparation with control parameters λ_{a}, enactment of protocol Λ, then work measurement during the conjugate protocol $\widetilde{\Lambda}$. Notice that the mean work beginning from a given microstate x, averaged over any distribution of microstates, is equal to the mean work of the ensemble of trajectories beginning from that distribution, e.g. $\left\langle \left\langle \mathcal{W} \right\rangle_{x;\widetilde{\Lambda}} \right\rangle_\Lambda = \left\langle \mathcal{W} \right\rangle_{\Lambda;\widetilde{\Lambda}}$. Similar relations do not hold for the higher order cumulants.

Relative entropy is not, in general, symmetric between two distributions, but if the distributions are similar, $q_x = p_x + \mathrm{d}p_x$, the asymmetry is small. A Taylor expansion of $\ln(q_x/p_x) = \ln(1 + \mathrm{d}p_x/p_x)$ about $\mathrm{d}p_x/p_x = 0$ gives, to leading order [9],

$$
D(p\|q) - D(q\|p) \approx \frac{1}{6}\sum_x \frac{(p_x - q_x)^3}{p_x^2} \ . \quad (10)
$$

Hence, we further assume that the asymmetry of the relative entropy is small in the near-equilibrium regime, $D(P_\Lambda\|P_{\lambda_{\mathrm{b}}}) \approx D(P_{\lambda_{\mathrm{b}}}\|P_\Lambda)$, and therefore that the sum of the higher-order excess work cumulants, $K_{\widetilde{\Lambda}}$, is directly related to the first excess work cumulant (the excess mean work) by

$$
K_{\widetilde{\Lambda}} \approx \frac{\beta}{2}\left(\left\langle \mathcal{W} \right\rangle_{\Lambda;\widetilde{\Lambda}} - \left\langle \mathcal{W} \right\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}} \right) \ . \quad (11)
$$

Substituting this approximation into Eq. 5 gives a more manageable expression for the near-equilibrium probabil-

ity distribution,

$$\ln \frac{P_\Lambda(x)}{P_{\lambda_{\mathrm{b}}}(x)} \approx -\beta\Big(\langle\mathcal{W}\rangle_{x;\widetilde{\Lambda}} - \langle\mathcal{W}\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}\Big)$$
$$+ \frac{\beta}{2}\Big(\langle\mathcal{W}\rangle_{\Lambda;\widetilde{\Lambda}} - \langle\mathcal{W}\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}\Big) . \qquad (12)$$

Inserting Eq. (12) into the equation relating free energy differences and relative entropy [Eq. (1)], it follows that the difference between the free energy of a nonequilibrium ensemble and the equivalent system in equilibrium, is approximately equal to minus one-half the excess mean time-reversed work:

$$\boxed{F_{\lambda_{\mathrm{a}},\Lambda} - F_{\lambda_{\mathrm{b}}} \approx -\frac{1}{2}\Big(\langle\mathcal{W}\rangle_{\Lambda;\widetilde{\Lambda}} - \langle\mathcal{W}\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}\Big) .} \qquad (13)$$

Here, finally, is our desired result. This free energy difference is readily measurable since it is approximately minus one-half the average work $\langle\mathcal{W}\rangle_{\Lambda;\widetilde{\Lambda}}$ when the system is prepared with protocol Λ starting from equilibrium at λ_{a} and then driven with the time-reversed protocol $\widetilde{\Lambda}$, less the average work $\langle\mathcal{W}\rangle_{\lambda_{\mathrm{b}};\widetilde{\Lambda}}$ when prepared in thermal equilibrium at λ_{b} and then driven with the time-reversed protocol $\widetilde{\Lambda}$.

Let us consider a simple system for which the above expressions are exact. A micron-sized bead is suspended in water by an optical laser trap. The trap is initially at rest, and then moves at a constant velocity, dragging the bead through the fluid. This system has been studied experimentally [10, 11] and can be modeled by a single particle undergoing diffusive Langevin dynamics on a moving, one dimensional harmonic potential. The pertinent properties of the model have been analyzed [12, 13]. Work distributions for a given initial particle position are Gaussian, with a mean work that depends linearly on the initial position of the particle relative to the center of the trap, and a position-independent variance. The third- and higher-order cumulants of the work are zero. Since the second and higher-order work cumulants are constant independent of the initial microstate, our first approximation, that $K_{x;\widetilde{\Lambda}}$ is independent of x [Eq. (7)], holds exactly. The equilibrium probability distributions are Gaussian, and, since the mean work is a linear function of position, it follows that the non-equilibrium probability distributions are also Gaussian, with the same variance, but shifted to a different mean relative to the equilibrium distribution. Consequently, our second approximation, Eq. (11), also holds exactly and our expressions for near-equilibrium probabilities and free energies are exact for this model at any driving rate.

Empirically, our expressions are good approximations across a more general class of systems. To demonstrate this, we explore a system for which the steady state probabilities, free energies, entropies and work distributions can be calculated exactly (within floating point accuracy). We simulate a particle diffusing over a periodic,
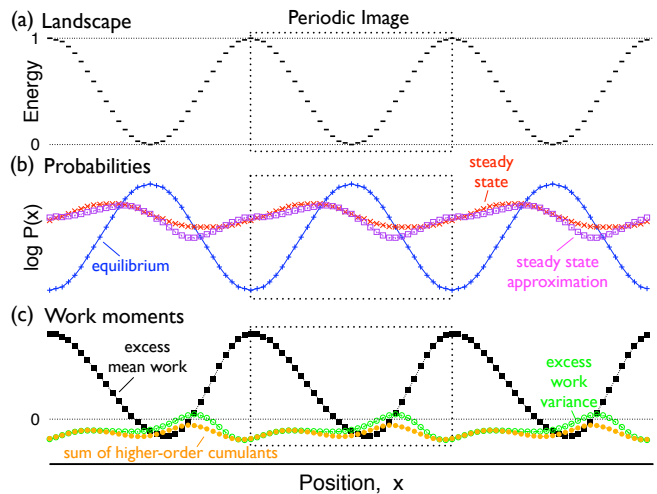

FIG. 1: (Color online) A simple driven system, amenable to numerical calculations. (a) Energy as a function of position. A single particle occupies a periodic, one-dimensional energy landscape. The position coordinate is discretized into $N_x = 32$ uniformly spaced positions per period. Energy is also discretized, $E(x) = \lfloor N_e(1 + \sin(2\pi x/N_x))/2\rfloor/N_e$ where x is position and $N_e = 64$ is the number of discrete energy bins. Results do not change appreciably with a finer discretization in either position or energy. (b) The system is initially in equilibrium with an external heat bath ($+$). At each discrete time step, the particle attempts to move one step left, one step right, or remain in the same location with equal probabilities, and the move is accepted according to the Metropolis criterion [16]. Every $1/v$ time steps, the energy surface shifts one position to the right. To ensure fully time-reversible dynamics, we simulate $1/2v$ time steps, shift the potential, and simulate another $1/2v$ time steps before examining the nonequilibrium properties of the system. All figures are drawn in the rest frame of the potential. Eventually the spatial distribution across a single periodic image converges to a non-equilibrium steady state ($\times$), approximated by Eq. (12) ($\square$). (c) The excess mean work ($\blacksquare$) varies markedly as a function of starting position, whereas the excess work variance ($\circ$) is relatively constant, and the sum of second- and higher-order cumulants $K_{x,\widetilde{\Lambda}}$ ($\bullet$) [Eq. (6)] is even moreso, implying compensating behavior. Displayed results are for $\beta = 4$ and $v = 1/4$. The reciprocal temperature β is reported in inverse units of the energy difference between top and bottom of the potential.

sinusoidal, one-dimensional energy landscape that moves at a fixed velocity (Fig. 1). Phase space and energy are discretized, hence all interesting properties of the system can be efficiently calculated using dynamic programming algorithms [14, 15]. See Fig. 1 for simulation details.

Figs. 2 and 3 demonstrate that for this model the steady state probability [Eq. (12)] and free energy [Eq. (13)] approximations are accurate given slowly shifting landscapes or high temperatures, and only diverge significantly from the exact results in strongly driven systems. We also find that qualitatively similar results arise for aperiodic potentials and for different potential sur-
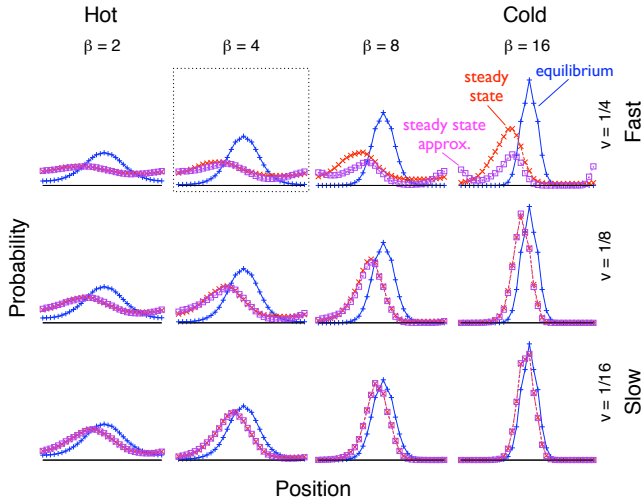
FIG. 2: (Color online) Equilibrium (+), steady state (×) and approximate steady state (□) [Eq. (12)] probability distributions, for the system described in Fig. 1, at various driving rates and temperatures. The quality of our approximate distributions, including overall normalization, deteriorates at low temperature and high driving velocity. The dotted box highlights the conditions shown in Fig. 1.
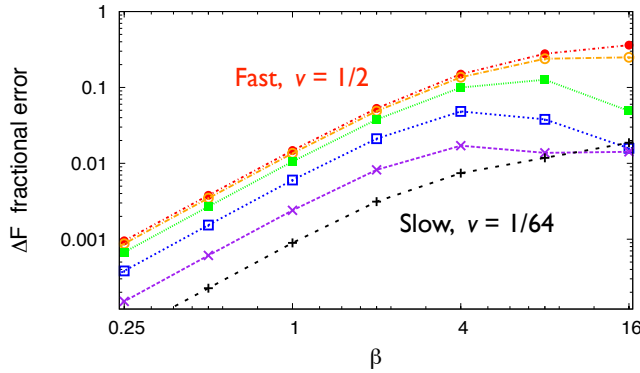

FIG. 3: (Color online) The approximate steady state free energy difference per periodic image, ΔF_{approx} [Eq. (13)], is very close to the exact steady state free energy difference $\Delta F_{\text{exact}} = F_{\lambda_a;\Lambda} - F_{\lambda_b}$, as shown by the fractional error $(\Delta F_{\text{exact}} - \Delta F_{\text{approx}})/\Delta F_{\text{exact}}$ being much less than unity. Each curve represents a set of simulations at a given velocity (ranging from fast ($v = 1/2$, top) to slow ($v = 1/64$, bottom)), with reciprocal temperature β increasing from left to right. Empirically for this system, ΔF_{approx} is always less than ΔF_{exact}, and the fractional error scales as β^2 for small β (large temperatures).

faces.

In this paper, we have developed a practical method for measuring free energies in the near-equilibrium regime, and our simulation results indicate that the approximate relation between free energy and excess mean time-reversed work is accurate a substantial distance from

equilibrium. Our analysis should be directly applicable to existing single-molecule experiments where the reverse protocol follows rapidly on the forward protocol, precluding equilibration [17]. We have concentrated on systems driven from equilibrium by a mechanical perturbation, but our relations could, in principle, be generalized to other situations, for example a system driven by a temperature gradient [18]. Verifying our approximations in more complex systems will require independent measurements of non-equilibrium free energies; one possible approach for simple fluids would be to computationally estimate entropies from multi-particle distribution functions [19].

[1] H. Qian, Phys. Rev. E **63**, 042103 (2001).
[2] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[3] D. J. Evans and D. J. Searles, Phys. Rev. E **52**, 5839 (1995).
[4] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1965).
[5] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, England, 1997).
[6] Y. Oono and M. Paniconi, Prog. Theor. Phys. Supp. **130**, 29 (1998).
[7] C. Jarzynski, Phys. Rev. E **73**, 046105 (2006).
[8] G. E. Crooks, Ph.D. thesis, University of California, Berkeley (1999).
[9] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge University Press, Cambridge, 2006).
[10] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002).
[11] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, Proc. Natl. Acad. Sci. U.S.A. **101**, 15038 (2004).
[12] O. Mazonka and C. Jarzynski (1999), cond-mat/991212.
[13] J. Horowitz and C. Jarzynski, Phys. Rev. E **79**, 021106 (2009).
[14] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, 2001), 2nd ed.
[15] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[16] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[17] H. Clausen-Schaumann, M. Rief, C. Tolksdorf, and H. E. Gaub, Biophys. J. **78**, 1997 (2000).
[18] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).
[19] A. Baranyai and D. J. Evans, Phys. Rev. A **40**, 3817 (1989).